Particle Flow Tuning of Fast Simulation for ILC.

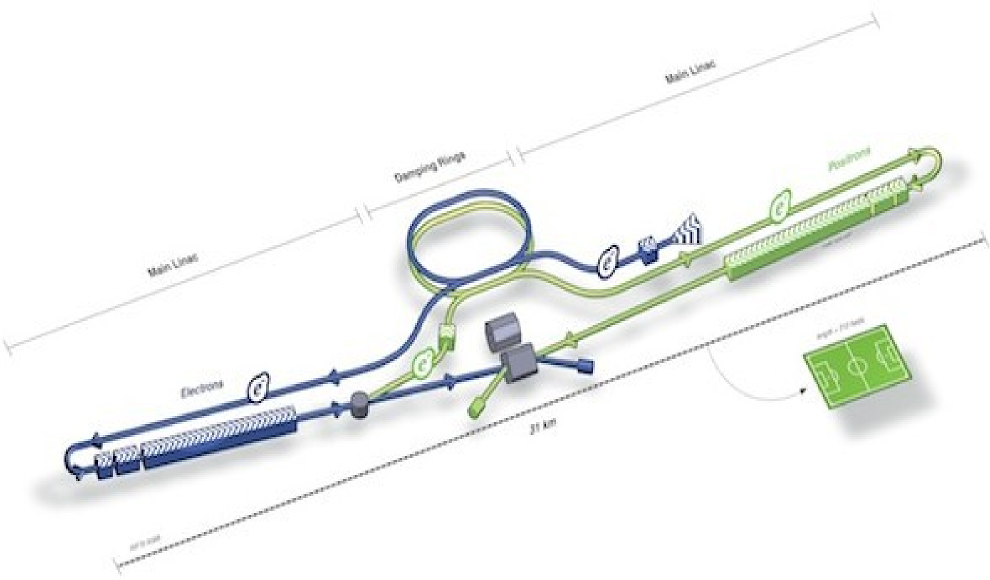
Eldwan Brianne, Mikael Berggren, Jenny List, Katja Krueger
FLC Group Meeting
14.04.14, Hamburg

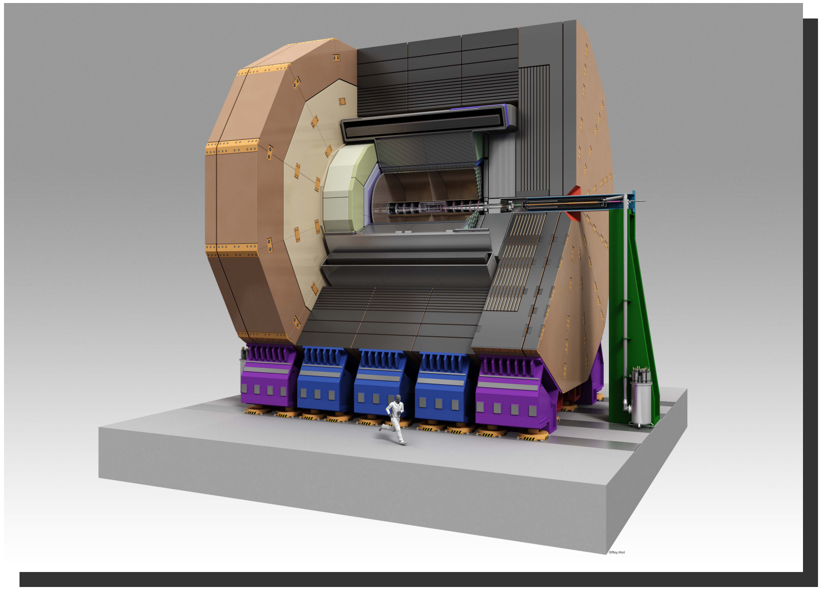
Outline

- ILC and ILD
  - Description and requirements
- A bit of physics : the amazing story of a hadron
  - Fragmentation and hadronization
- Particle Flow Concept
  - Association errors
- Fast Simulation and SGV
  - Comparison SGV without confusion & full simulation
  - Particle Flow emulation in SGV
- Remaining puzzle
  - Correlation in double counted and lost energy
  - First Ideas
- Conclusion & Outlook

The International Linear Collider & Detectors

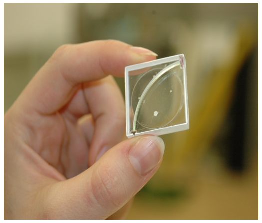
http://ilcild.org/





- Future e^+e^- collider @ 500 GeV (upgrade up to 1 TeV)
- Polarized beams : $P(e^+,e^-) = (0.3,-0.8)$
- Length : about 31 km (50 km @ 1 TeV)
- Motivation : High precision measurements and BSM searches

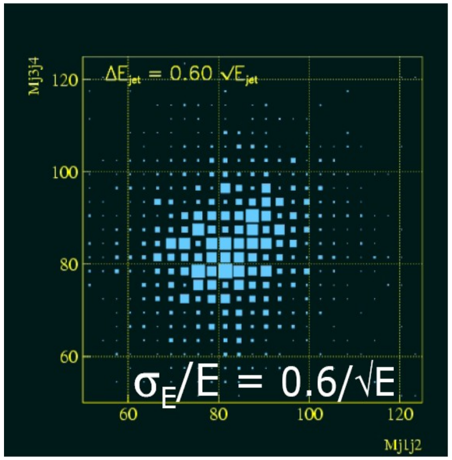
- Good vertexing, tracking resolution and high granular detectors
- The International Linear Detector (ILD)
  - Designed for Particle Flow
  - Compact detector and B = 3.5T
  - ECAL segmentation : 5*5 mm^2
  - HCAL tiles : 3*3 cm^2

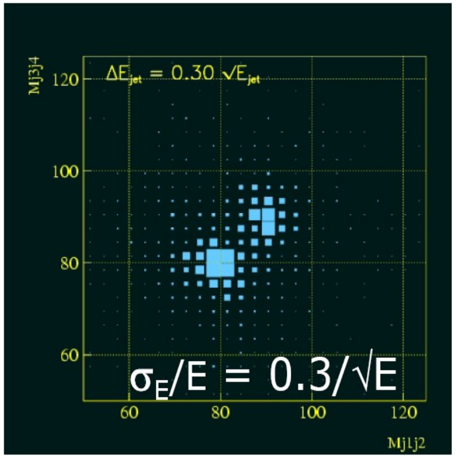


Detector requirements

- Detector performance requirements (high precision)
  - High momentum resolution : $\delta(1/p_T) \approx 5.10^{-5}$ GeV^{-1}
  - Jet energy resolution : ~ 3-4 %

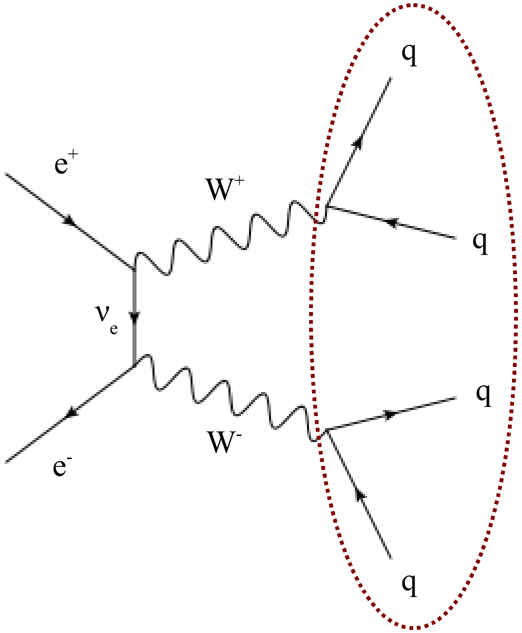
Needs a good understanding of the detector and the underlying physics





A bit of physics : the incredible story of a hadron

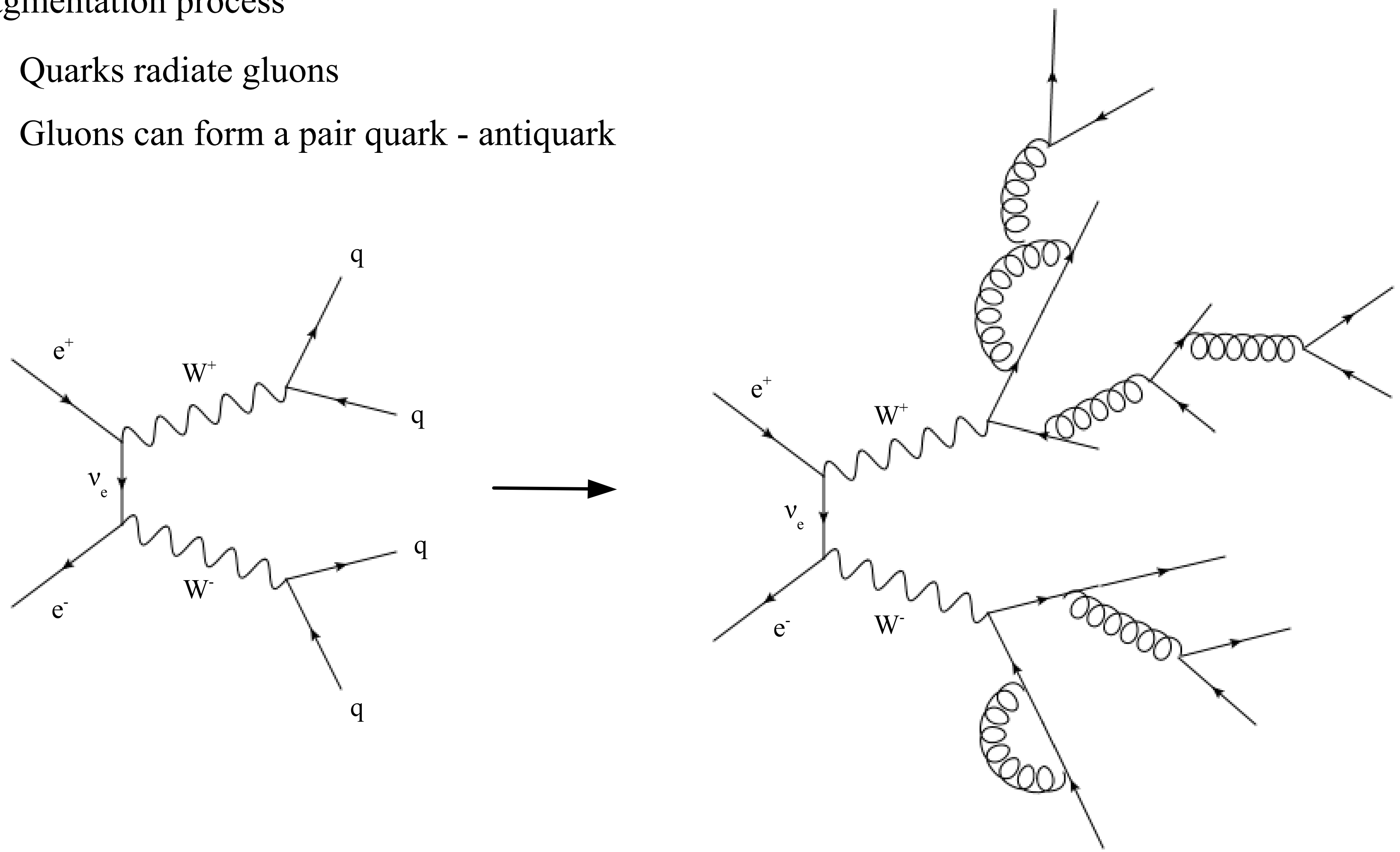
- Let's look at the process : $e^+e^- \rightarrow W^+W^- \rightarrow$ 4 jets @ $\sqrt{s}$ = 500 GeV



What is happening there ?
How do we get from quarks to hadrons
(what we actually detect in a calorimeter) ?

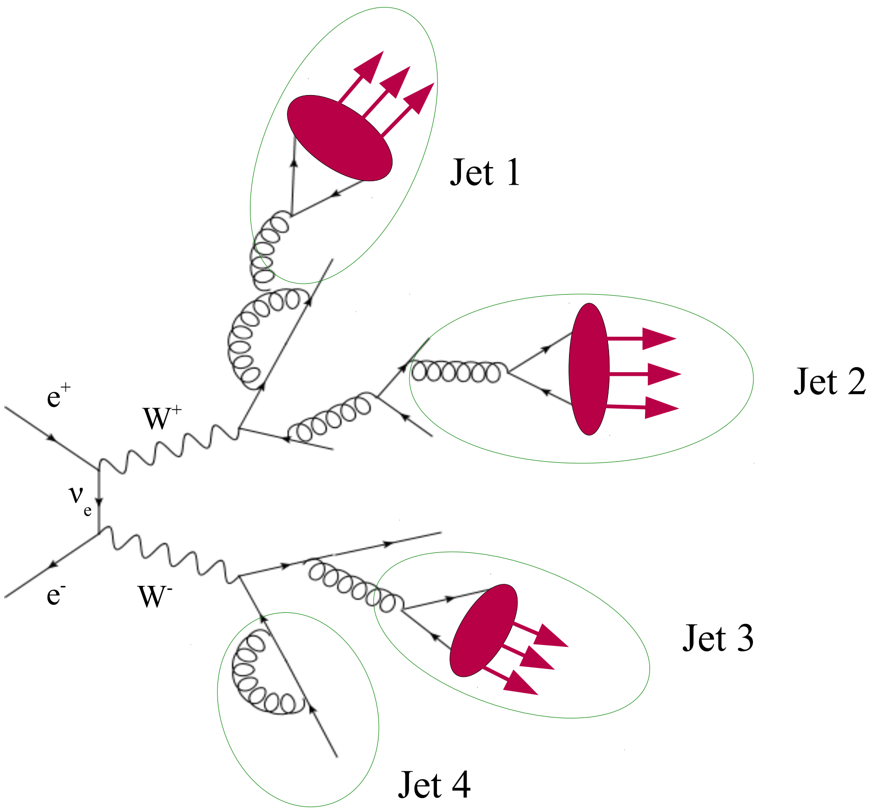
A bit of physics : the incredible story of a hadron

- Fragmentation process
  - Quarks radiate gluons
  - Gluons can form a pair quark - antiquark



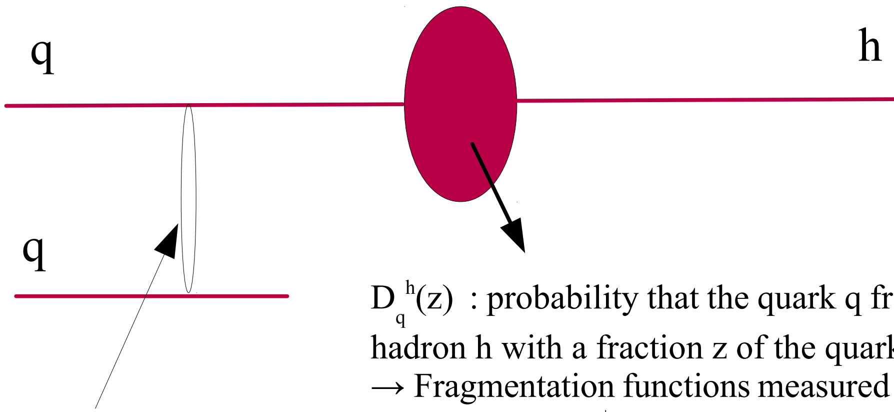
A bit of physics : the incredible story of a hadron

- This process continues until a certain energy scale → free quarks must then hadronize because of color confinement
- They form jets composed of many hadrons



- 2 models :
  - Cluster model
  - String model

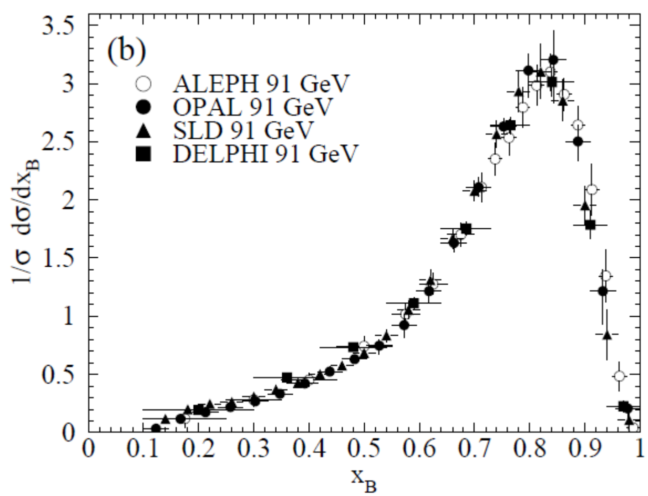
A bit of physics : the incredible story of a hadron

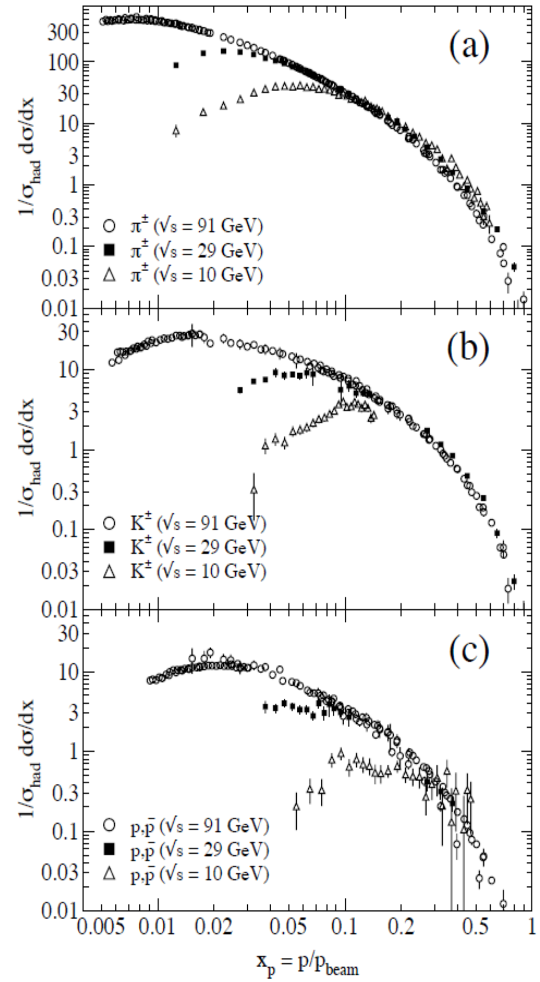


Strong color string
from gluon field

$D_q^h(z)$: probability that the quark q fragments into a hadron h with a fraction z of the quark momentum.
→ Fragmentation functions measured by several experiments in e^+e^-, pe^- and pp collisions : H1, ZEUS, ALEPH, OPAL...

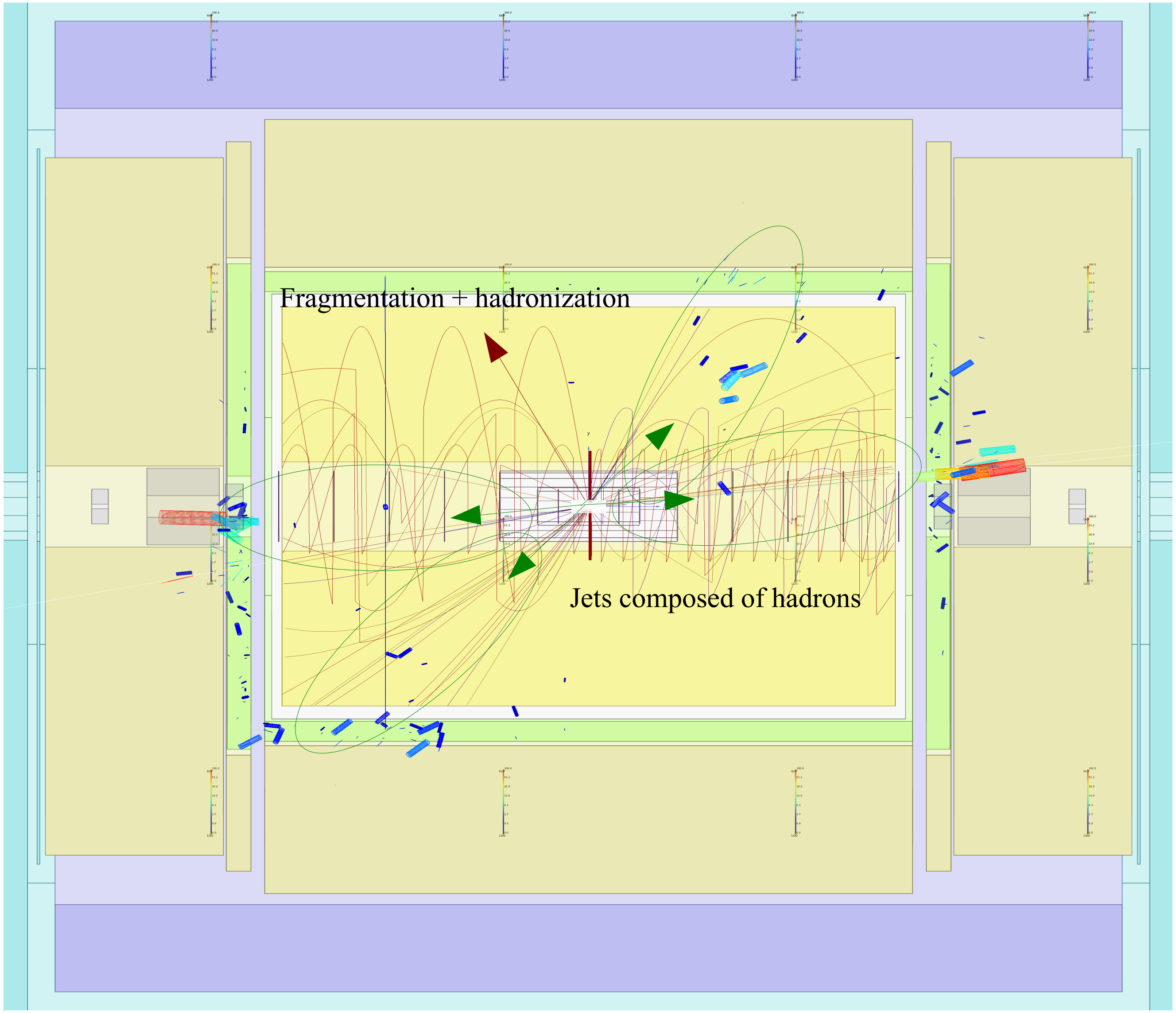
Bottom quarks - B hadrons





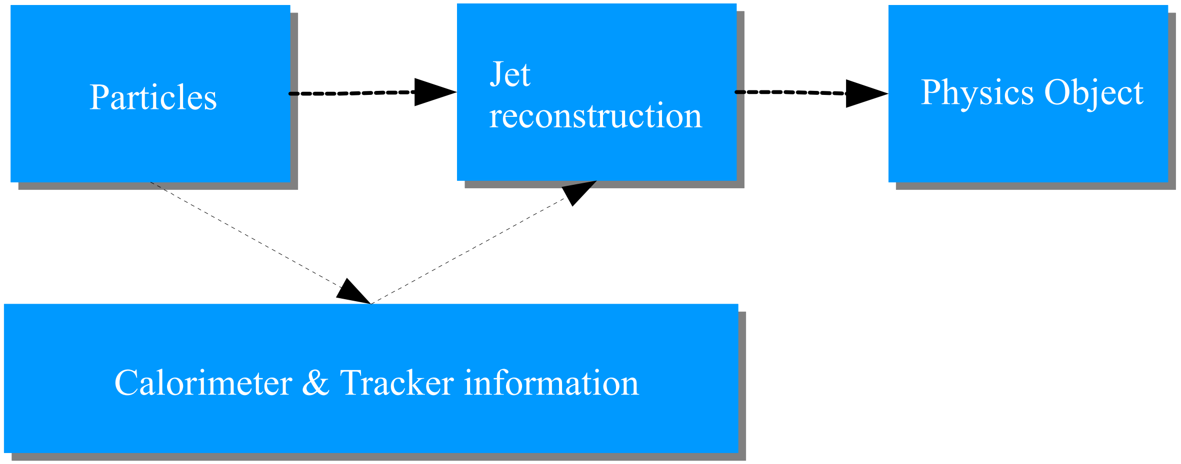
Fragmentation function
into pion, kaons and protons

A bit of physics : the incredible story of a hadron



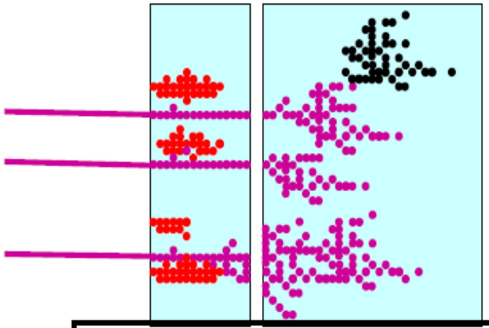
Getting back the physics objects

- How do we get back the physics objects : Higgs, W/Z bosons, top...?
  - Need to reconstruct jets : clustering algorithms, flavor tagging...
  - Unprecedented jet energy resolution => <u>Particle Flow needed</u>

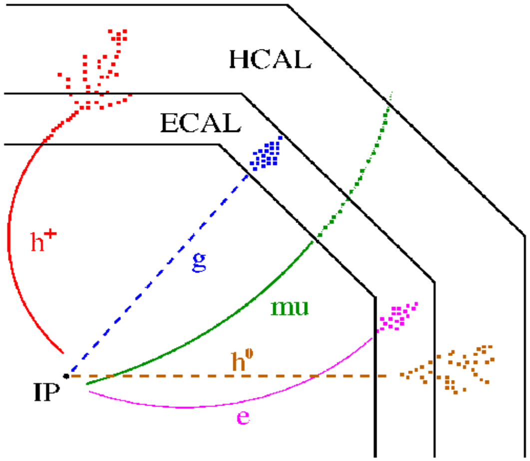


Particle Flow Concept

- Goal : Improve Jet energy resolution
- Principle :
  - identify and measure **every single particle in a jet** by making use of the information provided by all detector components
  - the jet energy is the <u>sum of the contributions of each individual particle</u>
- <u>To do so :</u>
  - Momentum measurement of charged particles in tracker
  - Energy of photons in ECAL
  - Energy of neutral hadrons in HCAL (only) – 10% fraction of the jet energy



$$E_{JET} = E_{TRACK} + E_{\gamma} + E_{n}$$

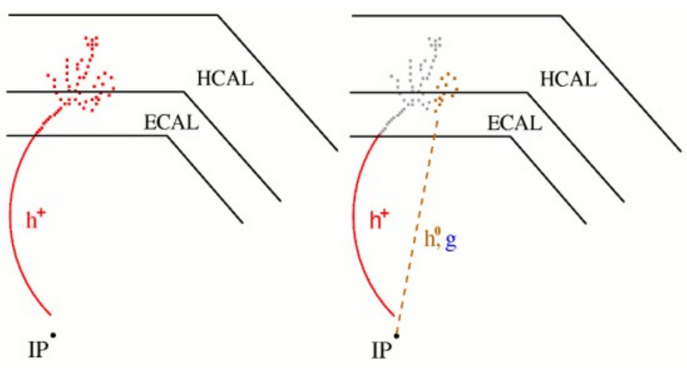


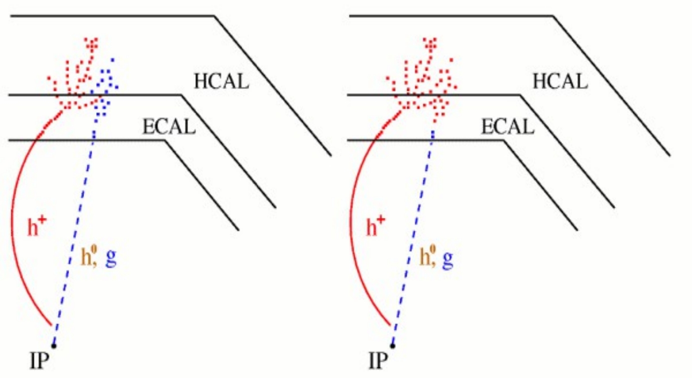
J. S. Marshall, M. A. Thomson arXiv:1308.4537v1

Association errors

- Association errors :
  - Cluster merging
  - Cluster splitting
  - Cluster - track mis-association
  - Other errors (charged to charged cluster, neutral to neutral cluster..) don't affect the total visible energy
- Confusion errors :
  - If a charged cluster is partially split to a neutral cluster → energy is double counted
  - If a neutral cluster is (partially) associated to a track → energy is lost

Charged cluster splitting

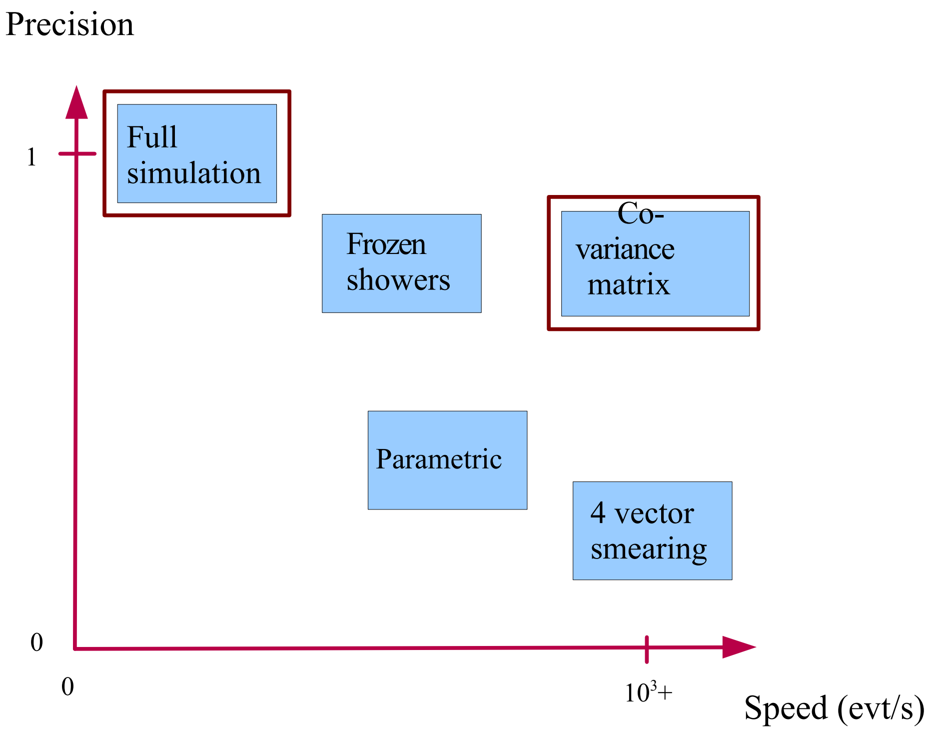




Neutral cluster merging and mis-association

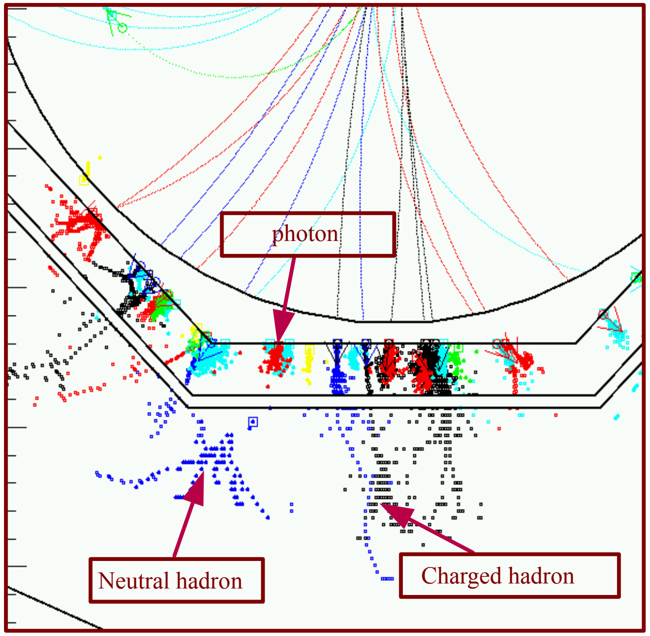
Fast Simulation

- Full Simulation : Geant4 – Mokka
  - Detector Optimization
  - Testbeam comparison
  - Drawback : O(mins)/events
- Can't be used for :
  - Physics studies with huge MC needs
  - SUSY searches (parameters scan), mass production, evaluation of systematics ($t\bar{t}$ events, W+jets, Z+jets)
  - Physics performance to different parameters (nominal energy, beam energy spectrum...)
- We need Fast simulation !!!
  - Simplified geometry
  - Validation with full sim & data



Fast simulation with particle flow detectors

- Fast simulation of a particle flow detector beyond 4-vector smeering :
  - Completely new approach and challenge
  - Never done before
  - Combining accuracy and speed for the needs of ILC
- Tool of choice : ....



SGV (Simulation a Grande Vitesse)

Advantages :

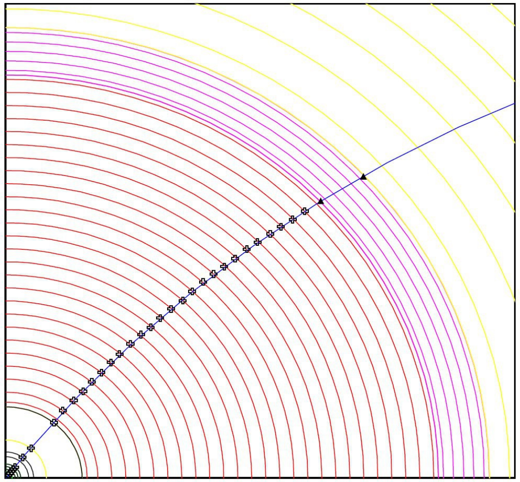
- Much faster than full simulation (~10^3 faster)
- Detector configuration can be changed easily
- Good agreement with full simulation

Tracking :

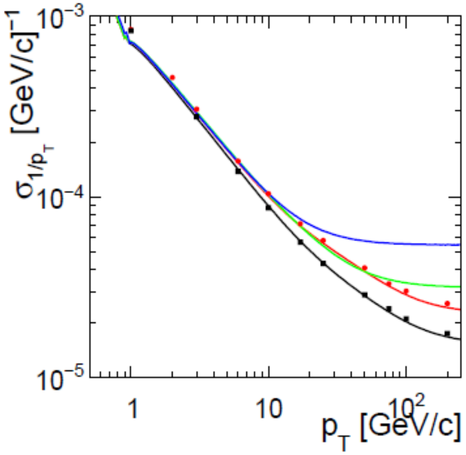
- Co-variance machine (inverse Kalman filter)
- Helix calculated by co-variance matrix at each layers, every parameters from first principles.
- Track resolution from detector geometry and point resolution → allows more flexibility

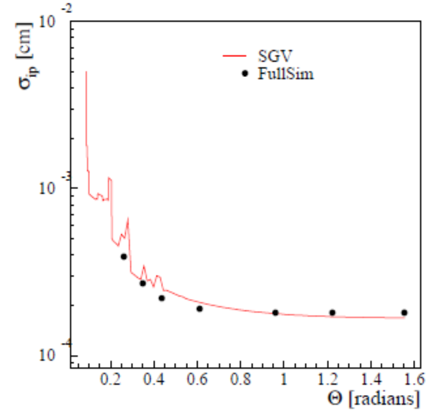
What about Calorimeters ?

M. Berggren, arXiv:1203.0217v1



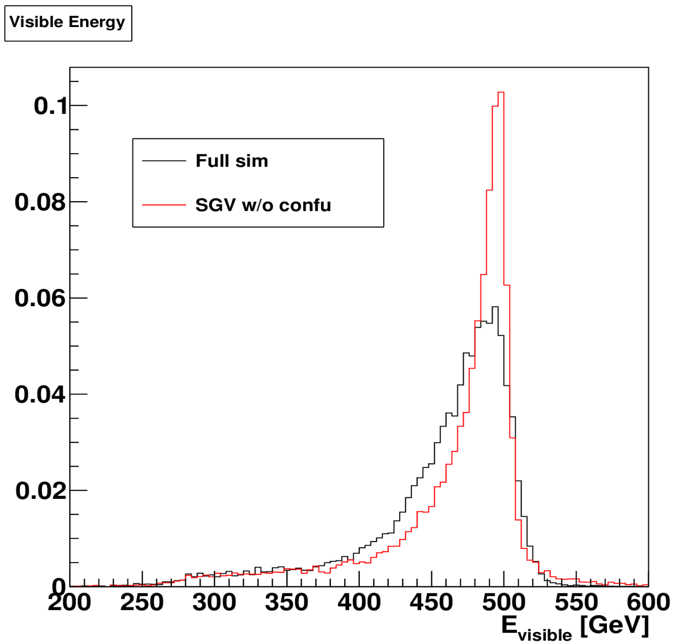
Tracking in SGV

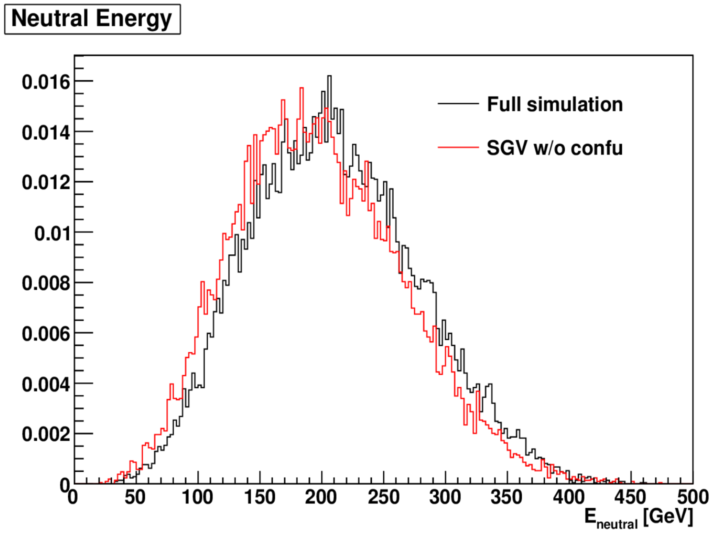


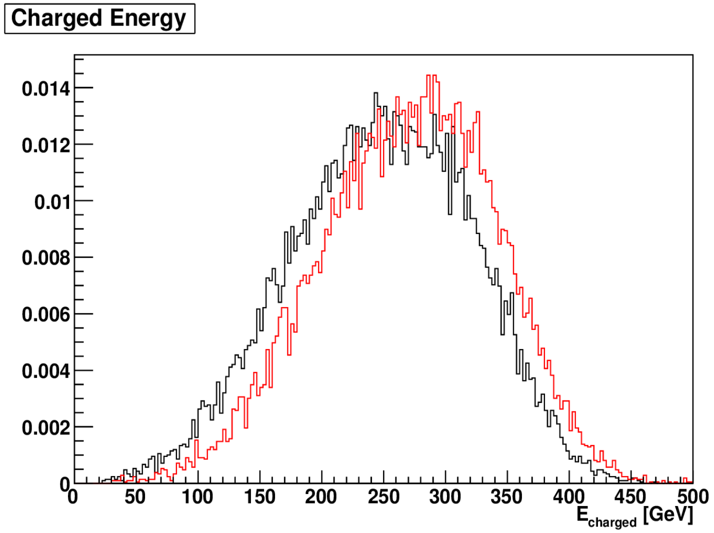


Comparison SGV w/o confusion / full simulation

- Process : $e^+e^- \rightarrow W^+W^- \rightarrow 4$ jets
- Total Energy : peak too narrow
- Neutral and Charged energy : resolution in agreement but shift in energy!
- Need particle flow emulation!

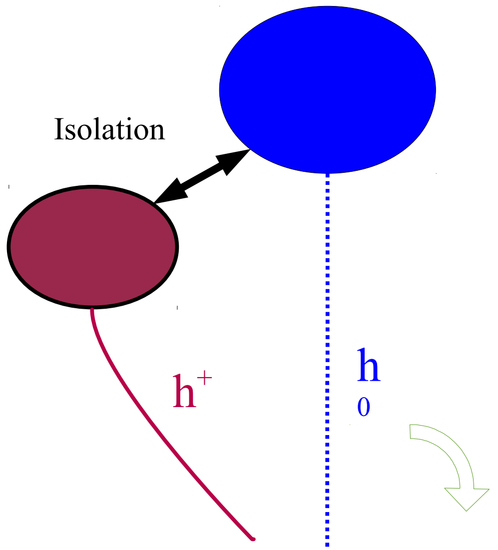


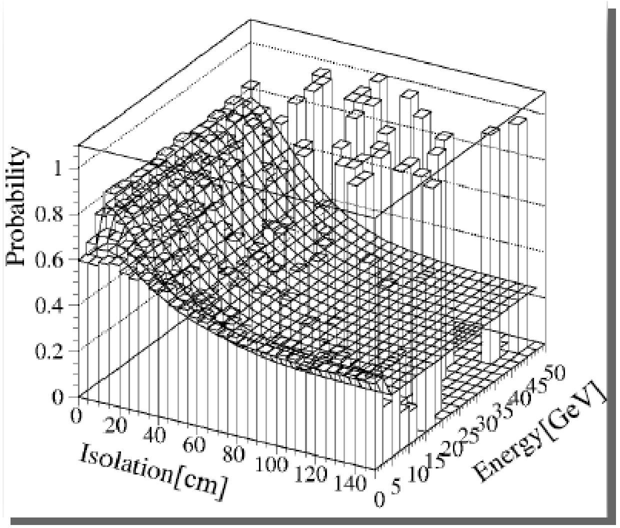




Particle Flow in SGV

- Algorithm :
  - First, clusters are made from simulated showers (calculated by a function) of each particle
  - Determine parameters like type, distance & energy of the closest neighbor
  - Simulate confusion by calculating a probability of splitting or merging a cluster depending on the isolation, the energy of the cluster and the type (EM or hadronic)
  - Same output as full simulation in order to compare !
- What has been achieved :
  - Tuning of SGV to PandoraPFA by parametrizing the association errors → depending on energy, isolation of the cluster, type of the particle, detector region and if it is EM or hadronic showers (in total 28 parameters)

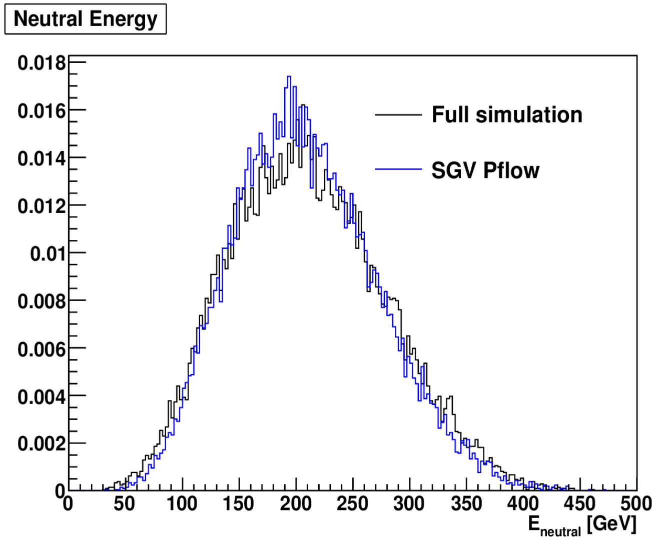


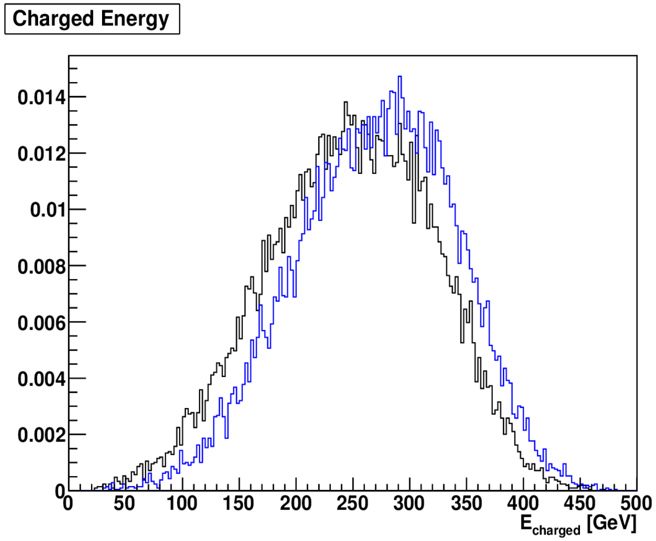


M. Berggren, arXiv:1203.0217v1

Comparison SGV Particle Flow / full simulation

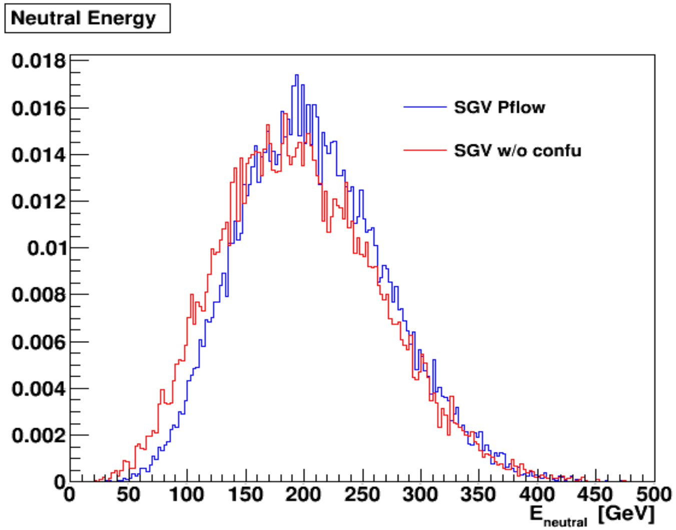
- With current Particle Flow tuning
  - Much improved agreement for neutrals
  - Still shift in energy for charged particles but the resolution is fine (5% difference) – Almost no change compared to no confusion

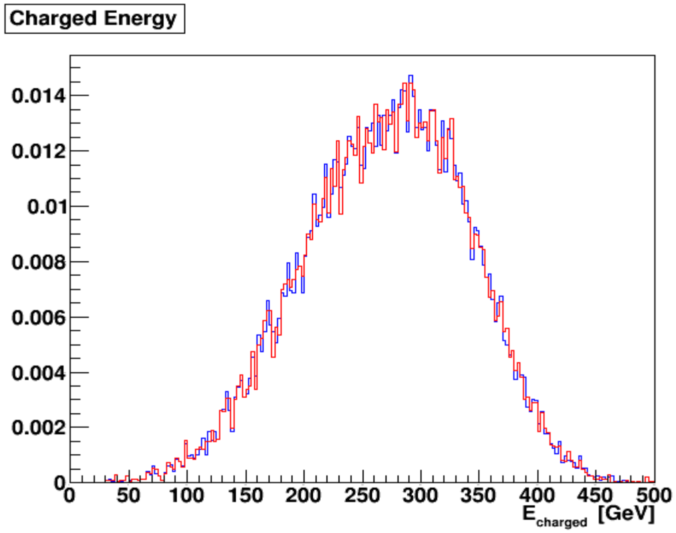




Comparison SGV Particle Flow / full simulation

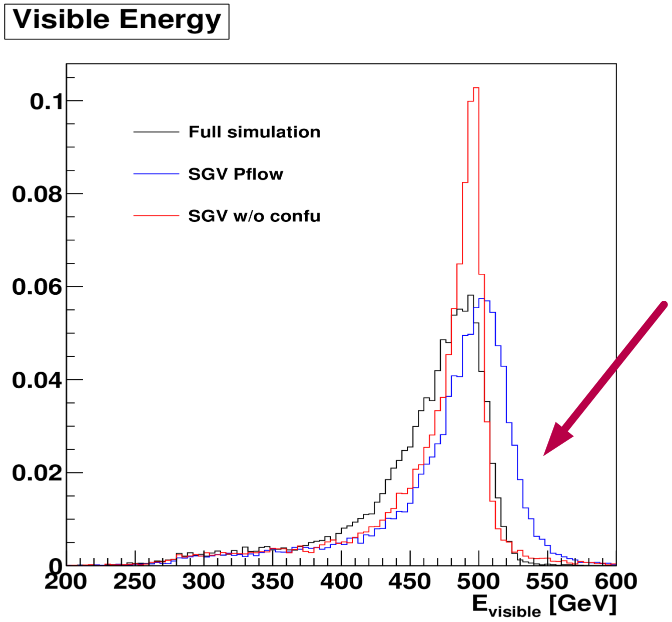
- With current Particle Flow tuning
  - Much improved agreement for neutrals
  - Still shift in energy for charged particles but the resolution is fine (5% difference) – Almost no change compared to no confusion

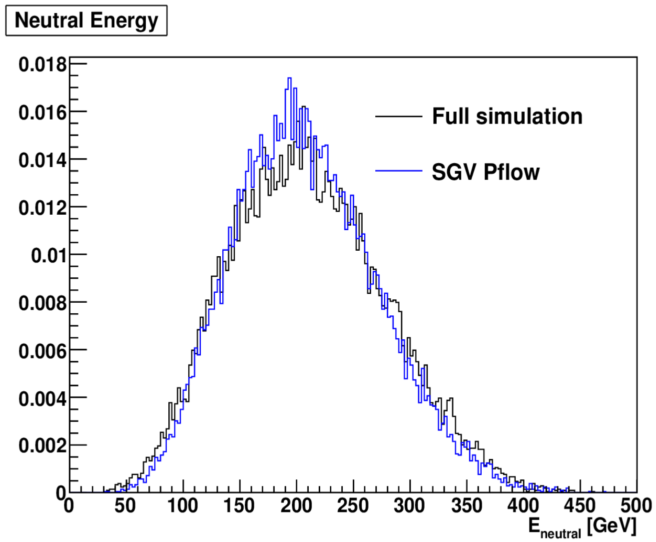


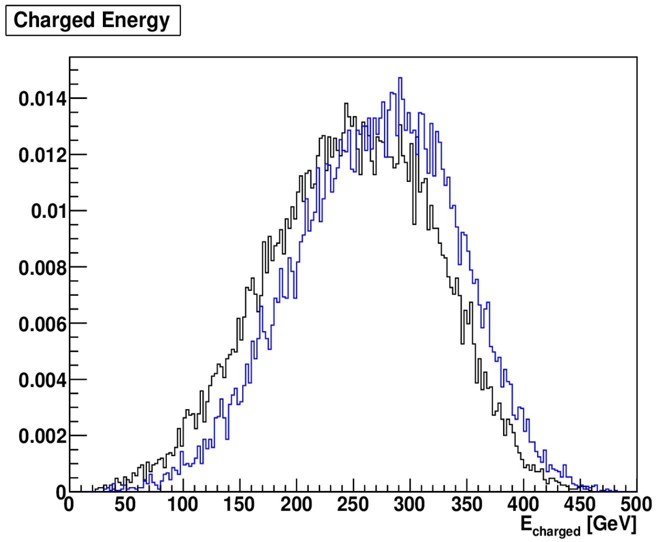


Comparison SGV Particle Flow / full simulation

- With current Particle Flow tuning
  - Much improved agreement for neutrals
  - Still shift in energy for charged particles but the resolution is fine (5% difference) – Almost no change compared to no confusion
  - Induce a tail on the total visible energy

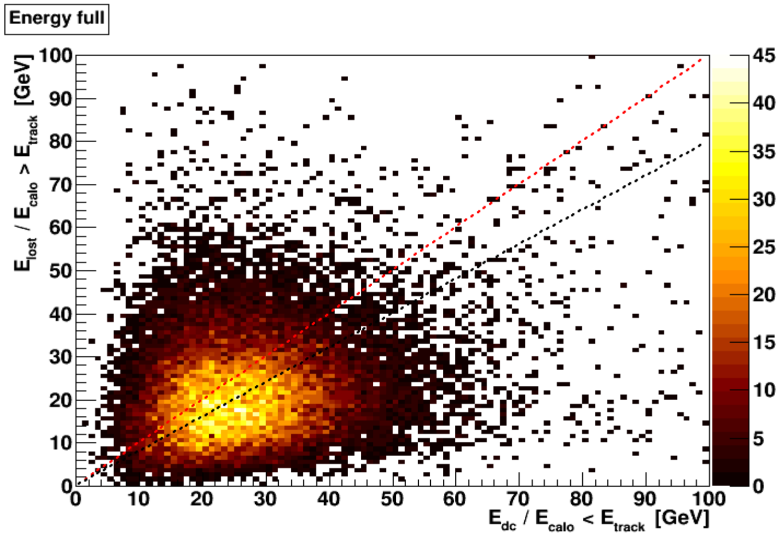


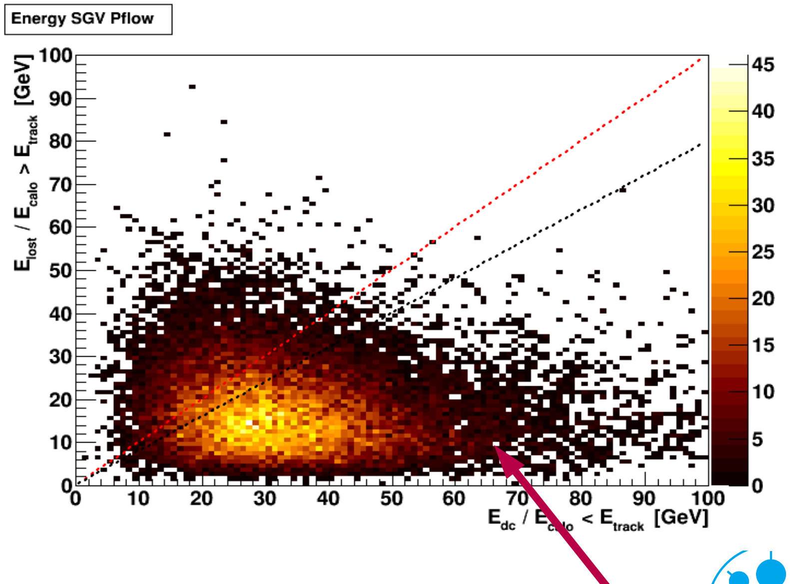




Comparison SGV Particle Flow / full simulation

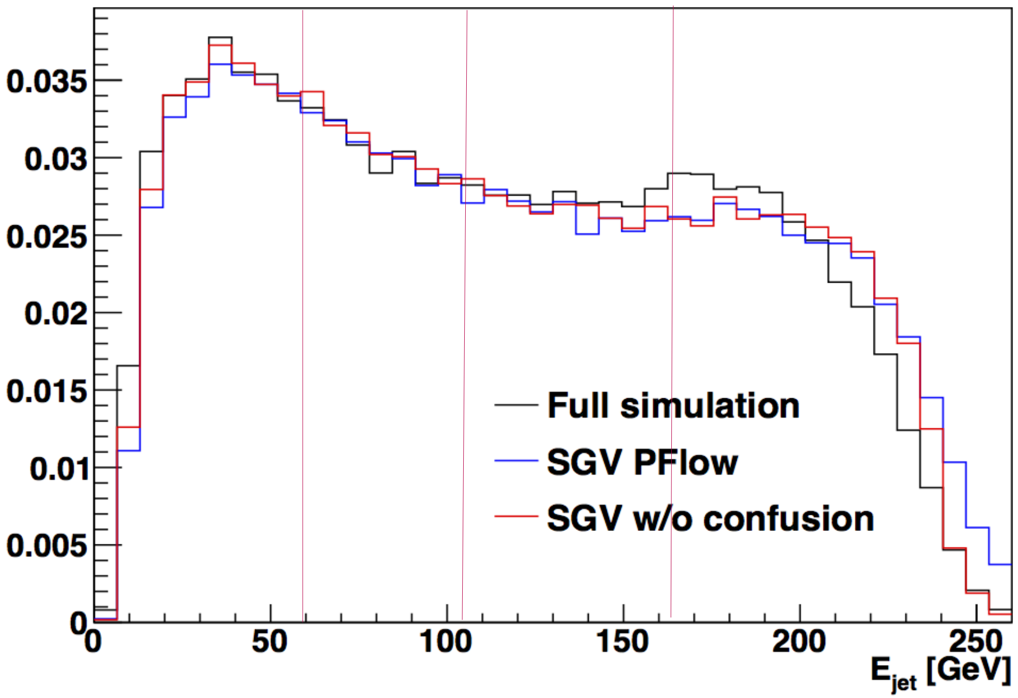
- Define double counted and lost energy (only for charged particles/event):
  - E_{dc} if $E_{calo} < E_{track}$
  - E_{lost} if $E_{calo} > E_{track}$
- With current Particle Flow tuning
  - More double counting & less lost energy in SGV
  - Correlation E_{dc}/E_{lost} needs to be tuned for SGV





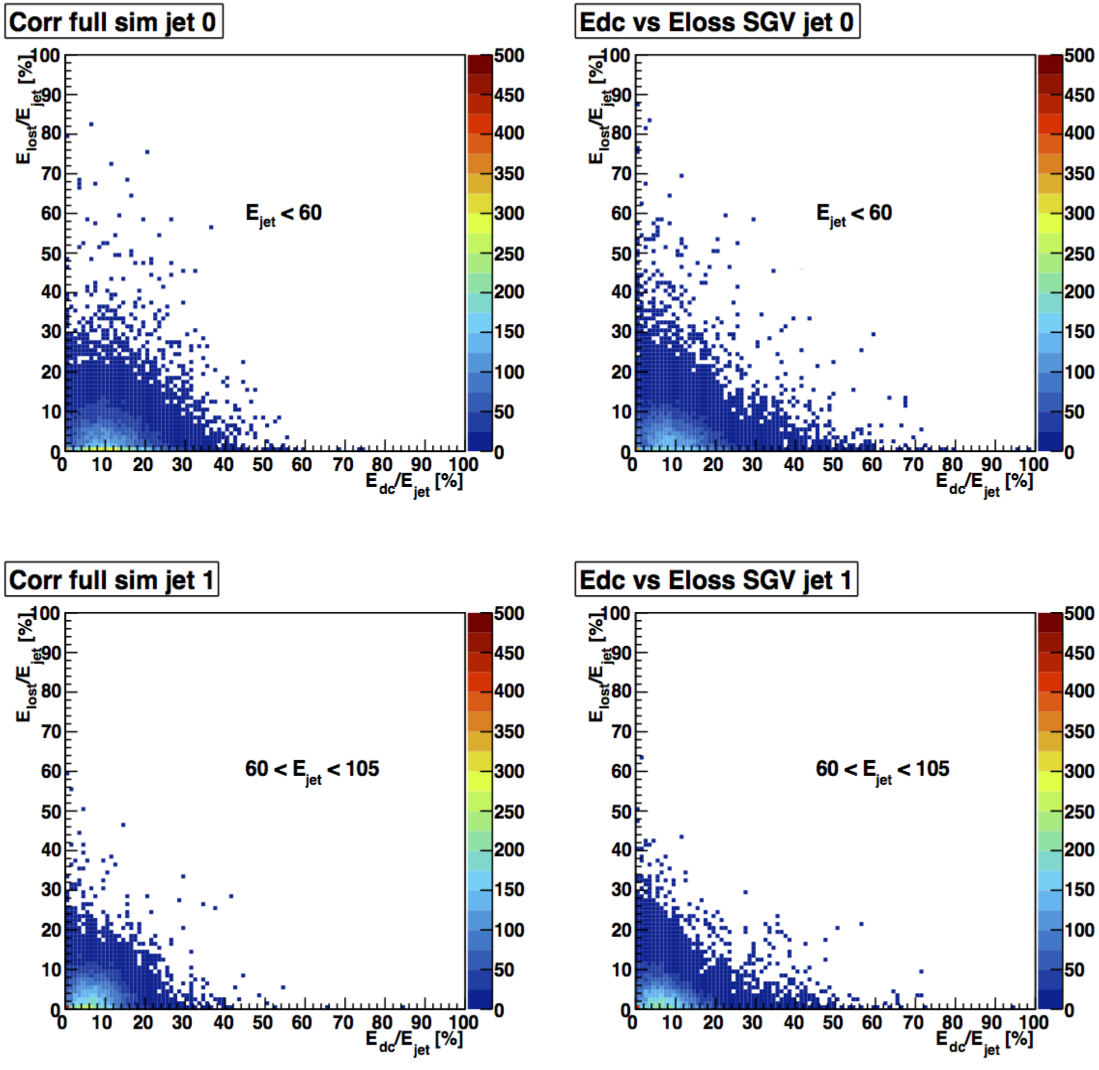
Looking more into details

- Look the correlation at a jet level for charged particle inside the jet. Double counted and lost energy defined per jet.
- Done in function of the jet energy (4 bins) :
  - $E_{jet} < 60$ GeV
  - 60 GeV $< E_{jet} <$ 105 GeV
  - 105 GeV $< E_{jet} <$ 165 GeV
  - $E_{jet} > 165$ GeV



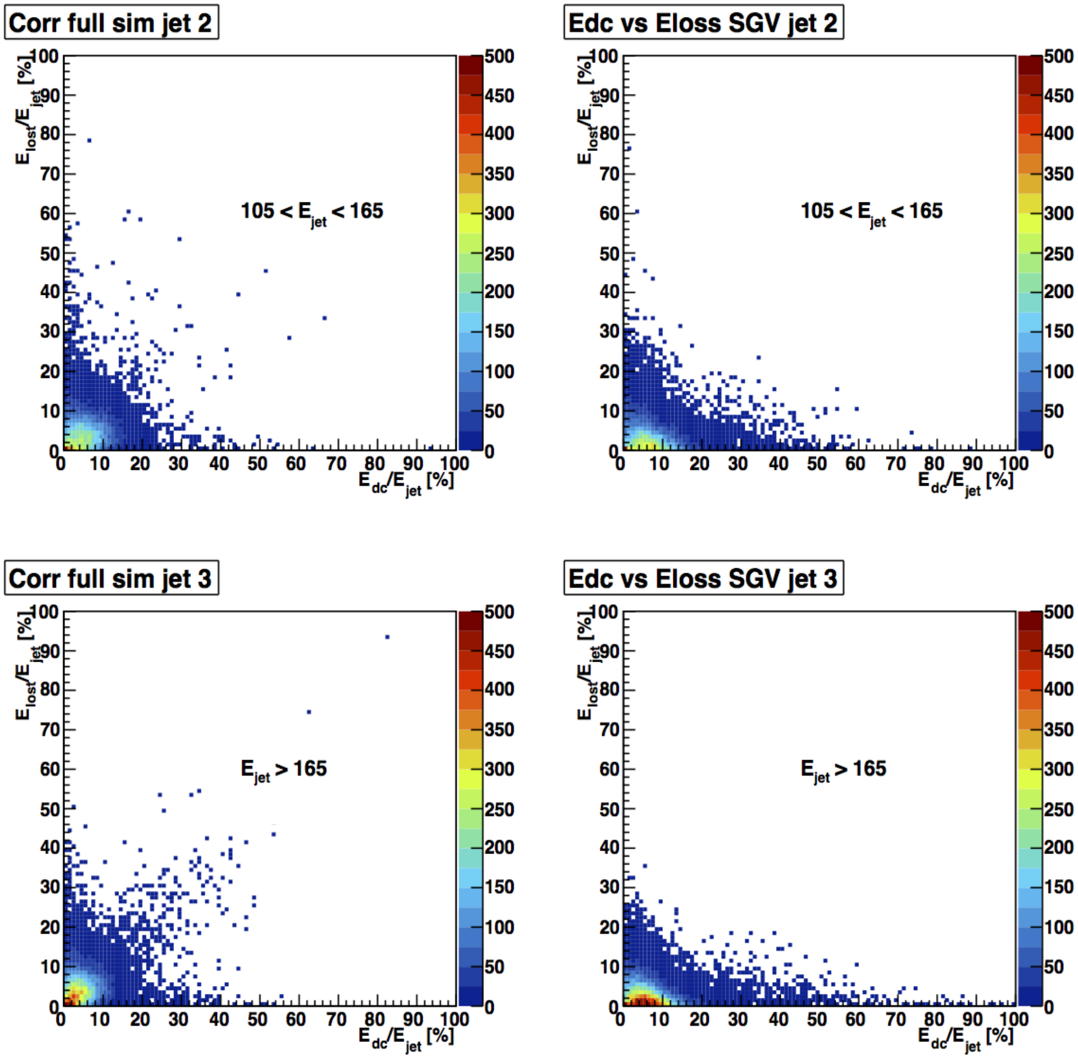
Energy correlation per jet

- Seems to be ok for low energy jets.



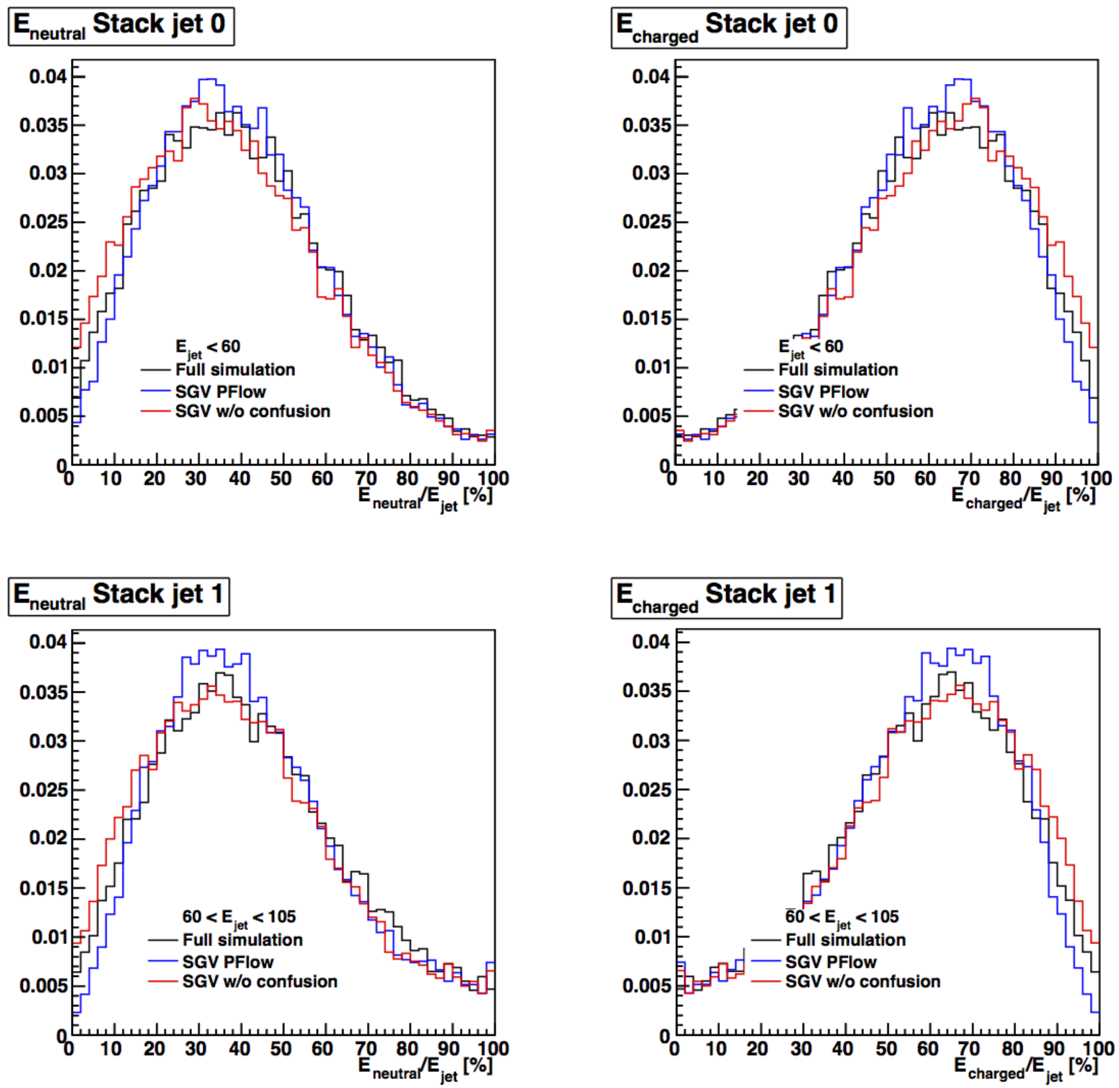
Energy correlation per jet

- Full simulation shows a significant correlation for high energy jets which is not present in SGV.
- Tail for double counted Energy in SGV.



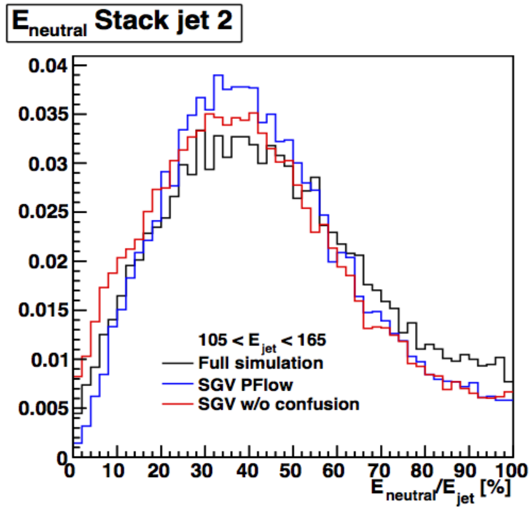
Neutral & Charged energy per jet

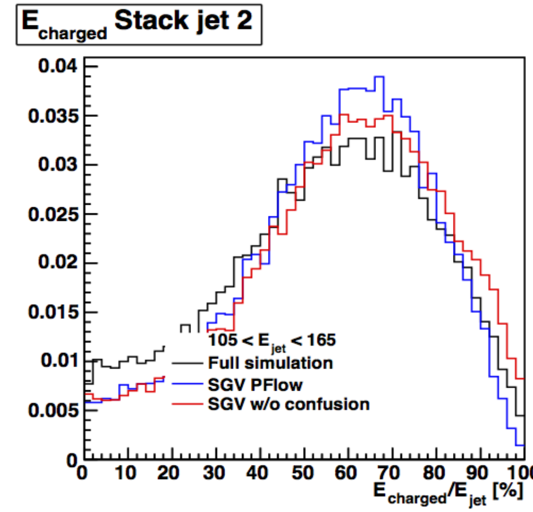
- Rather good agreement for low energetic jets
- Neutral & Charged energy are mirrors of each others

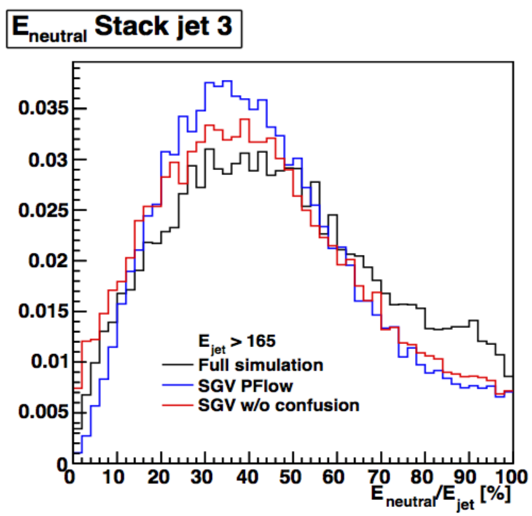


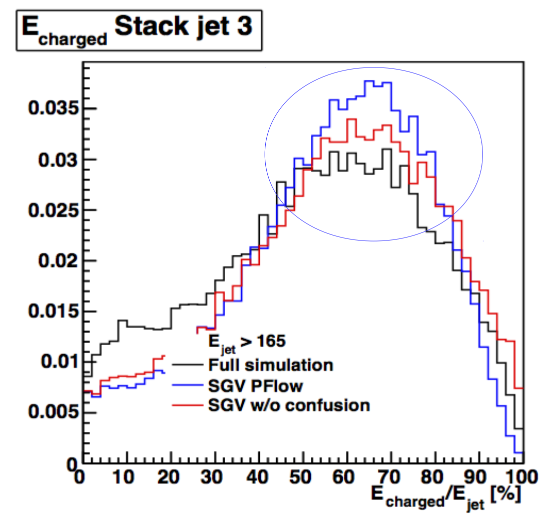
Neutral & Charged energy per jet

- Particle Flow parametrization seems increase the energy in the 20-50% region → should do it over 50% region
- Need to reduce the charged energy in 20-50% region







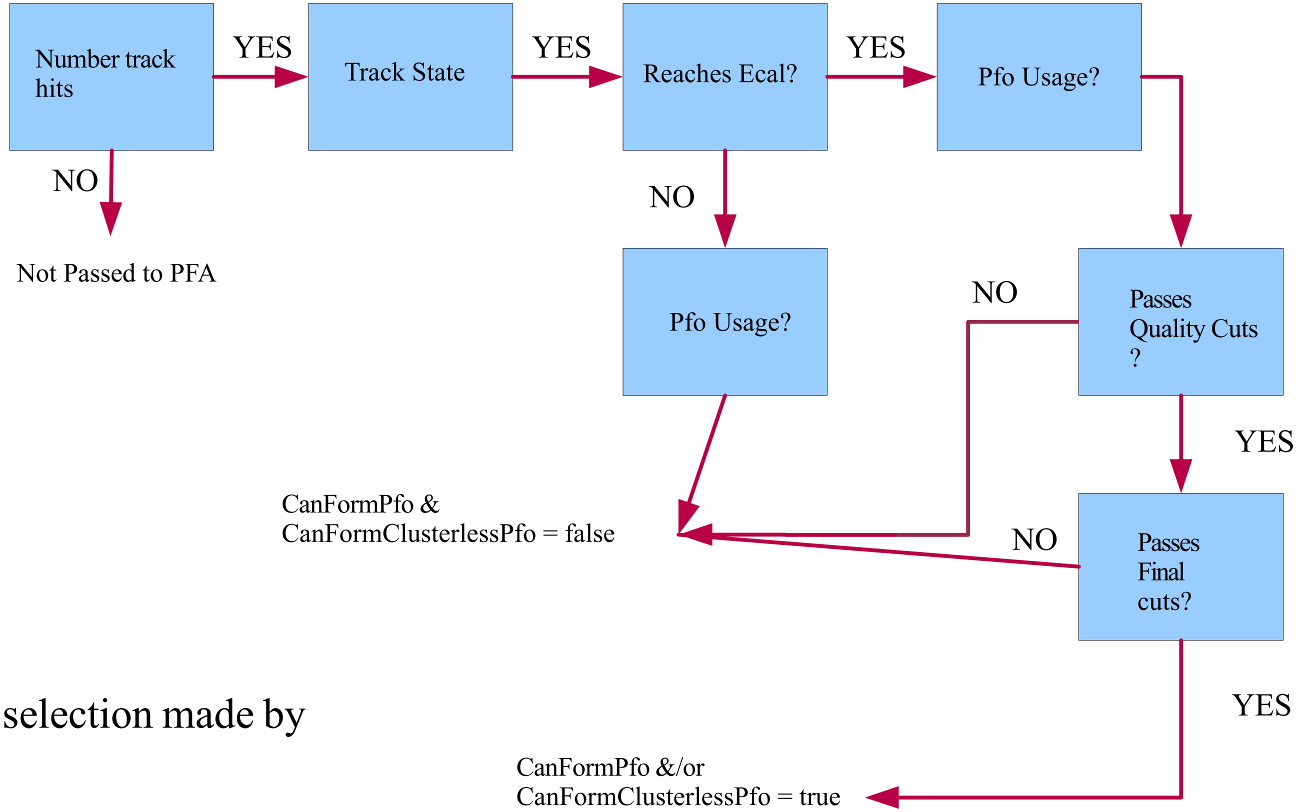


First Ideas

- First few ideas to do :
  - In case of complete split $\rightarrow$ reject the track (look into it first)
  - Relax E/p criterion
  - Looping several times over the cluster list
  - Checking the jet angular dependence

- Redo the plots and look at the changes

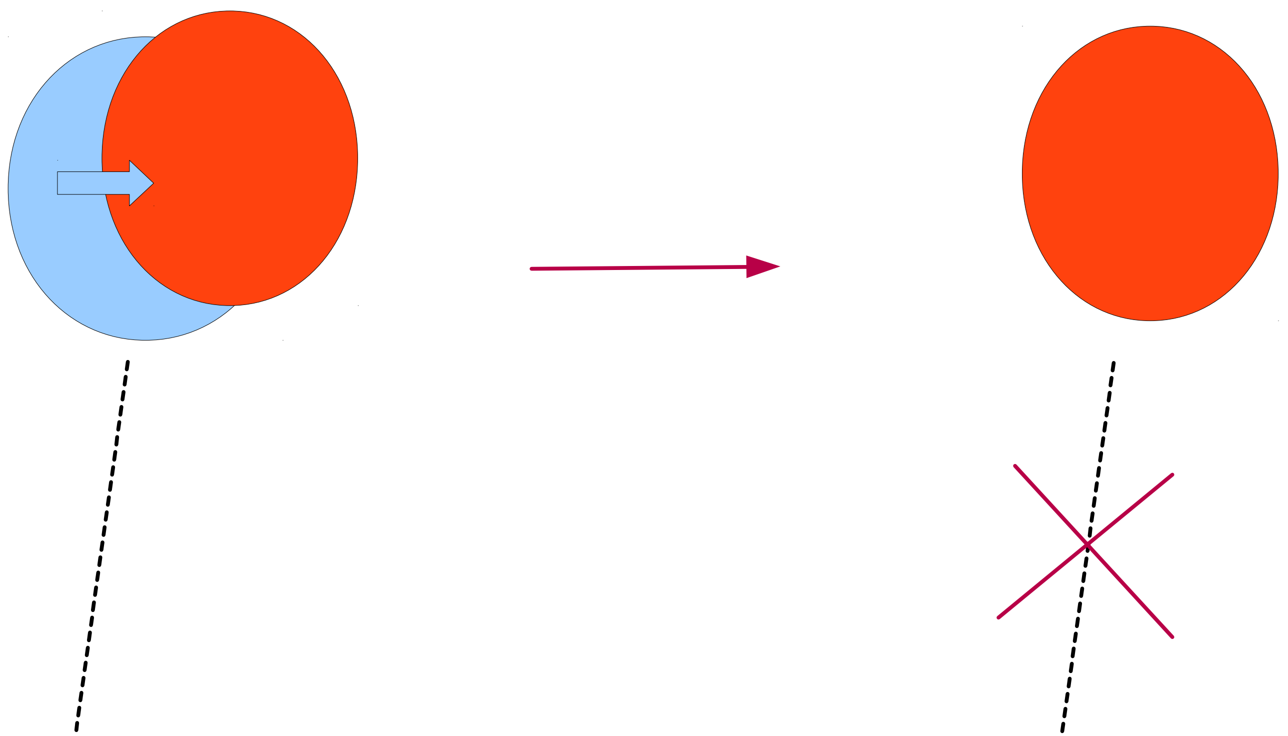
First idea : Track selection in Pandora



- Understand the track selection made by Pandora
- Categorization based on track parameters : d0, z0, r_{min} and track energy
- Drop tracks in case of complete split, taking into account Pandora criterions.

First Idea : Current status in SGV

- Idea :
  - Complete split → the energy of the charged cluster is then assigned as a neutral cluster, in addition the track is kept.
  - Need to drop the track in this case → reduction of charged energy



Conclusion & Outlook

Conclusion

- Fast simulation with particle flow calorimeter is a new challenge
- SGV is already close to what Pandora is doing in full simulation
- Remaining puzzle : correlation between loss and double counted energy

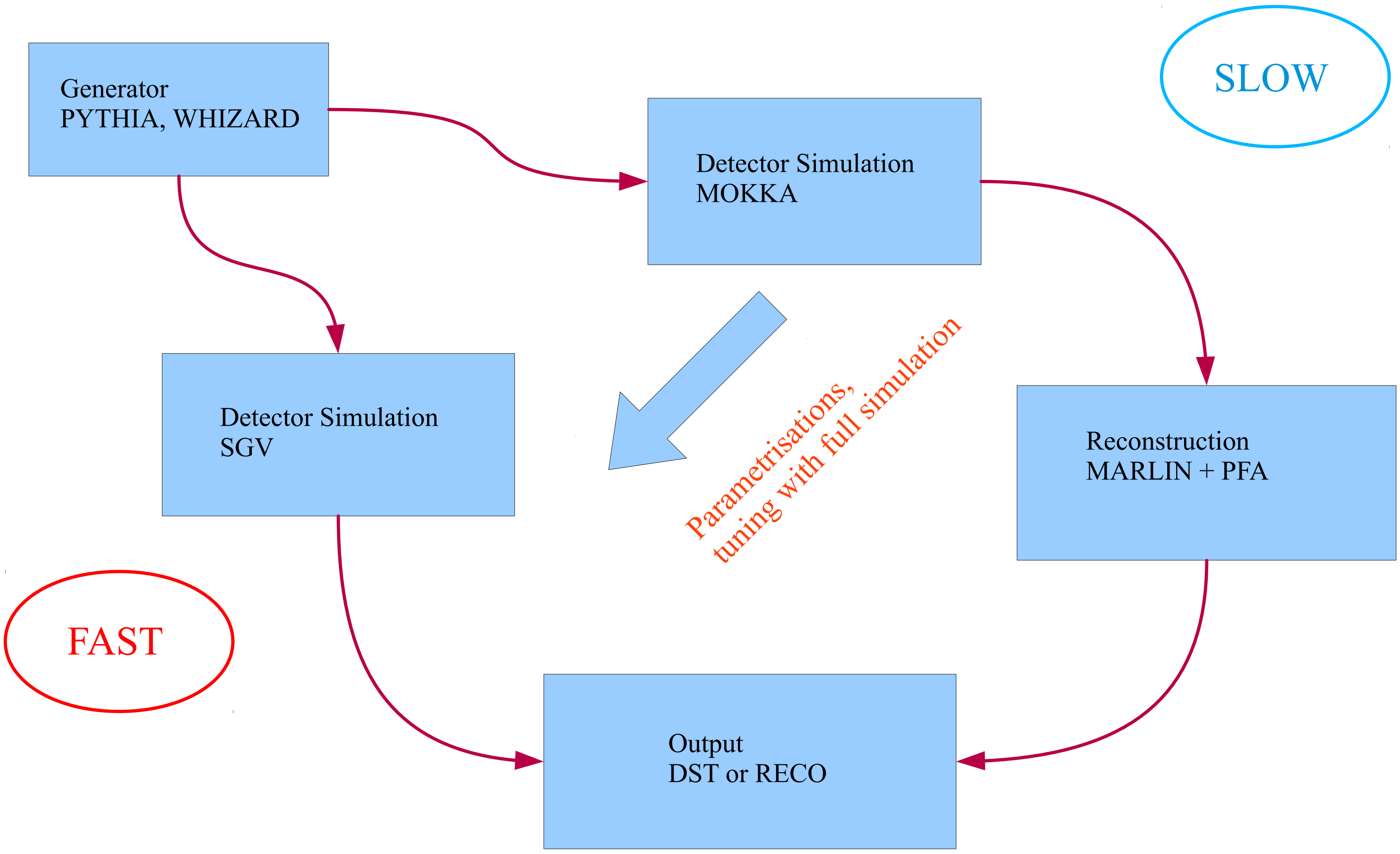
Outlook

- Place for improving the merging, splitting algorithm
- Maybe doing energy flow in dense region
- Need to understand the correlation between lost energy and double counted energy in full simulation in order to tune SGV
- Applying track rejection in case of complete split & look changes

Thanks for your attention !!

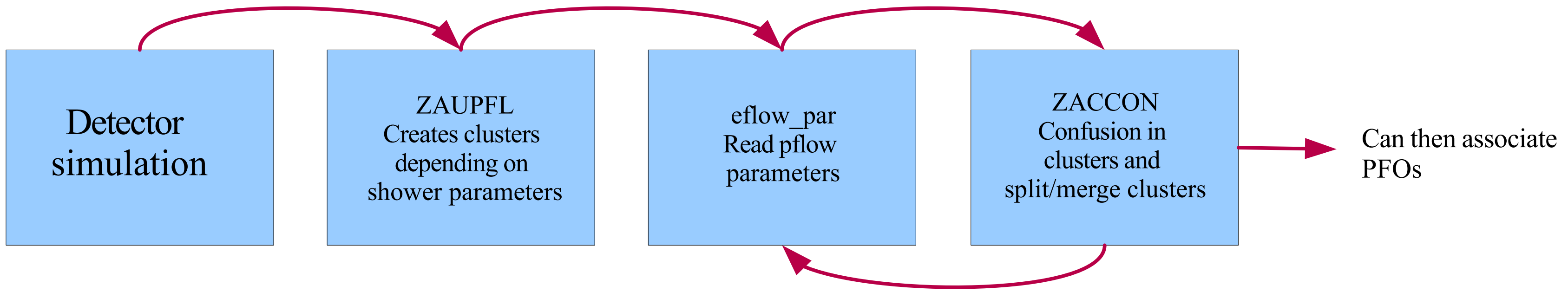
Backup Slides

Backup : Simulation Framework



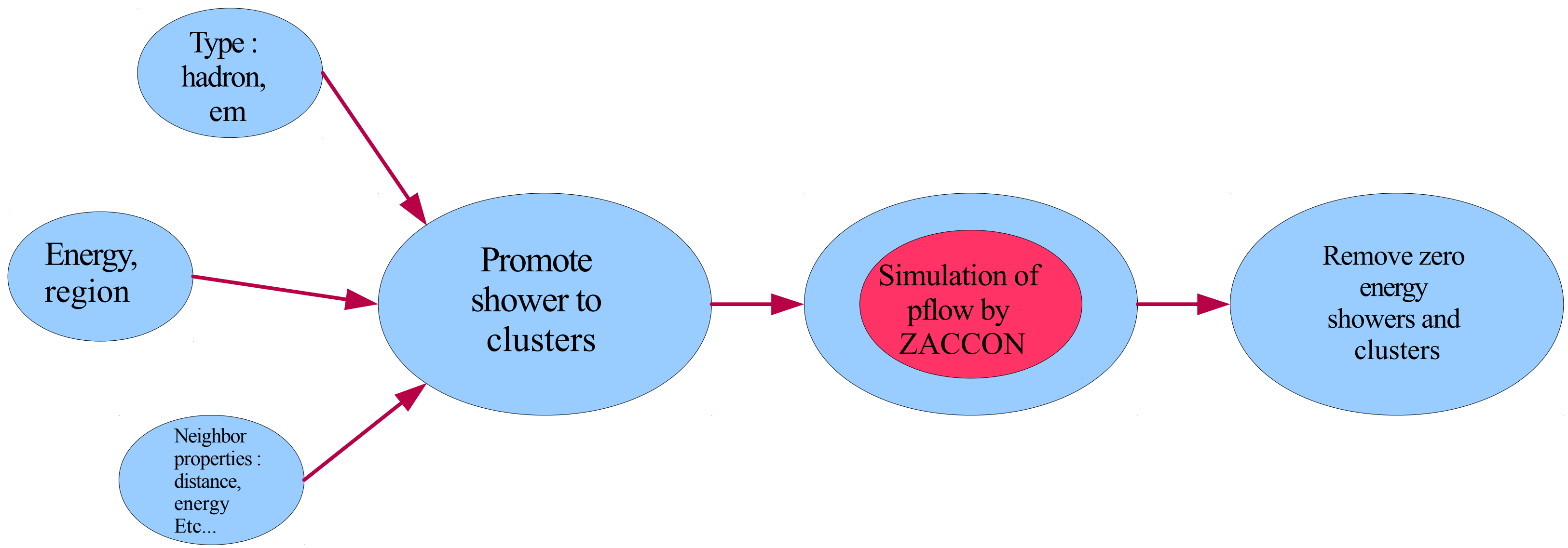
Backup : How is Pandora implemented in SGV?

- Particle Flow can be enabled in SGV (by changing the steering parameter for the Calo to PFL instead of PERF)
- Pflow is implemented by a routine named ZAUPFL which simulate Pandora (shower shapes, splitting and merging into clusters)
- Then a routine ZACCON is called to simulate confusion in clusters (can be dummy or not)
- In ZACCON, a routine eflow_par is called, declare pflow parameters (depending on dc or loss energy) and a do a routine to read them



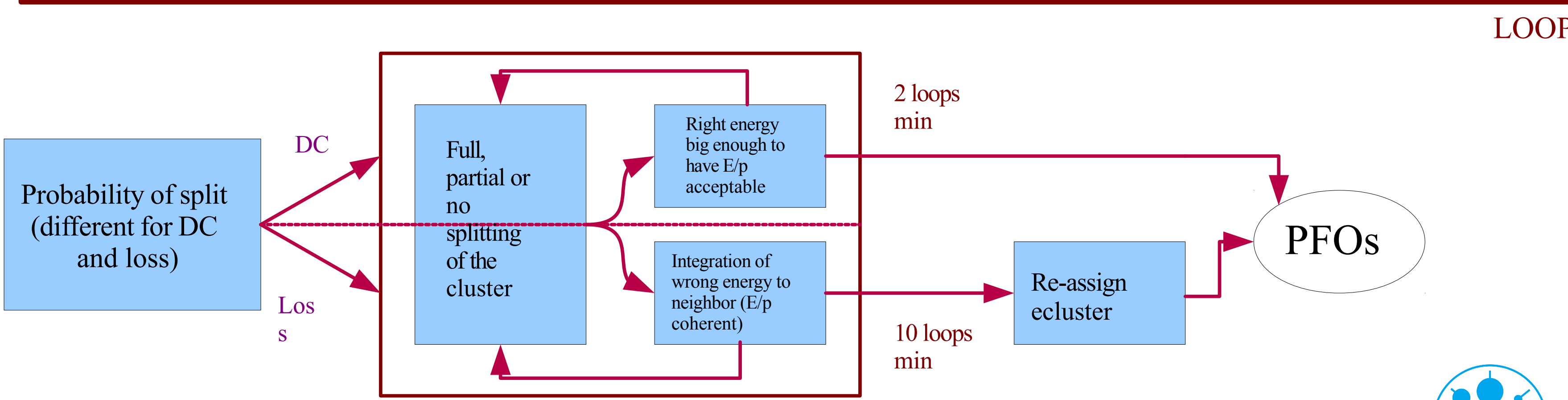
Backup : Routine ZAUPFL

- Used to promote showers to clusters
- Figure out parameter for each shower : type, closest neighbor of the same type, energy, barrel or forward calorimeter, other closed clusters
- Simulate particle flow with inputs : cluster properties, closest neighbor and distance. Output the energy kept by the particle and lost energy. Calling the ZACCON routine.
- Remove zero energy clusters and showers after confusion routine.



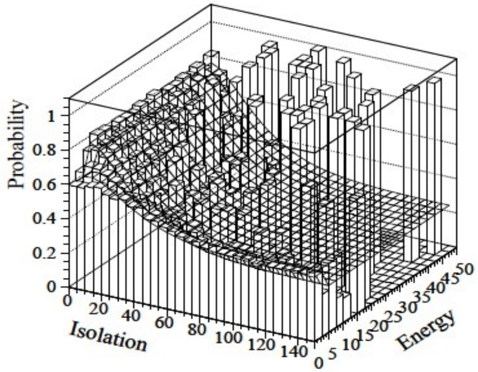
Backup : Routine ZACCON

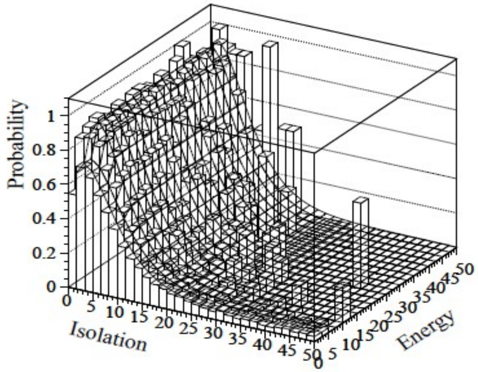
- Routine to simulate confusion
- Calculate a probability (function using parameters and few pseudo-random numbers) of split in case of dc or loss energy
  - If complete split, assign to a neutral cluster with full energy
  - If partial, calculate the em fraction of right energy in the cluster (depending on distance to the closest neighbor, energy and random number and assign rest to wrong energy.
  - Loss case :
    - Look if wrong energy can be integrated with the neighbor respecting a coherent E/p (check 10 loops) – then re-assign cluster energy.
    - If no neighbor, no splitting
  - DC case :
    - Look if right energy has still a coherent E/p (check 2 loops), re-loop if not



Backup : Eflow_par module

- Module in order to declare and read which parameters to use for particle flow depending on the type of particle, and the case of double counting or loss of energy.
- Read parametrization files : em_dc_5.dat, em_loss_5.dat, had_dc_5.dat, had_loss_5.dat
- Parameters :
  - spvsdnp, spvsdnpkc2, spnorm, zprob
  - tfit2, avefrac_e, avefrac_d
  - avefrac_norm
- Theses parameters correspond to the fit parameters of 2D confusion probability functions

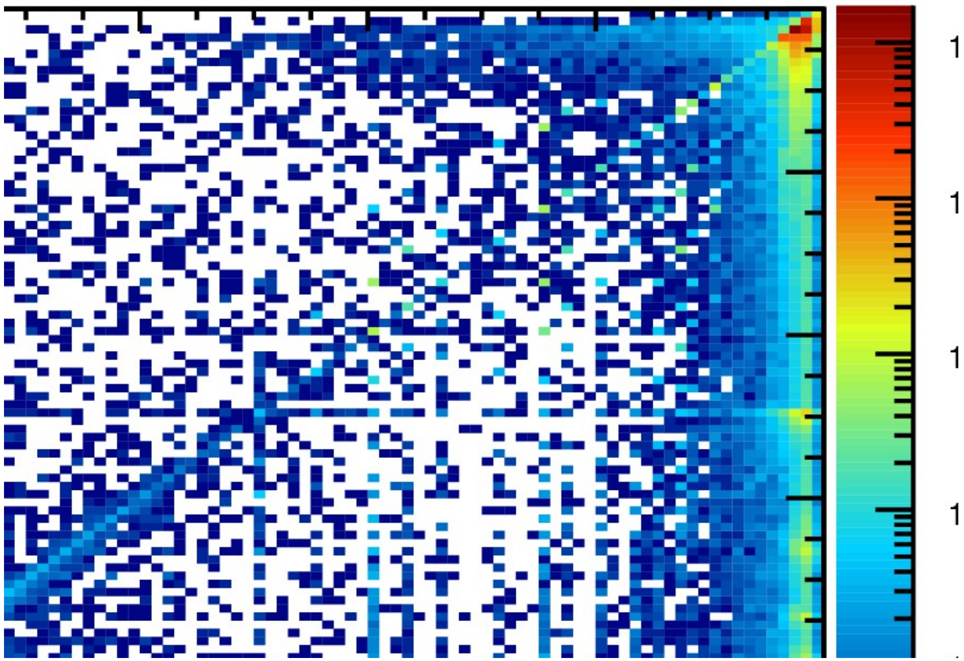




Backup : Parametrizing Particle Flow Performance

- Now we need to go more in detail
- Use of RecoMCTruthLinker in full relations
- Need to understand how Pandora is working in full simulation

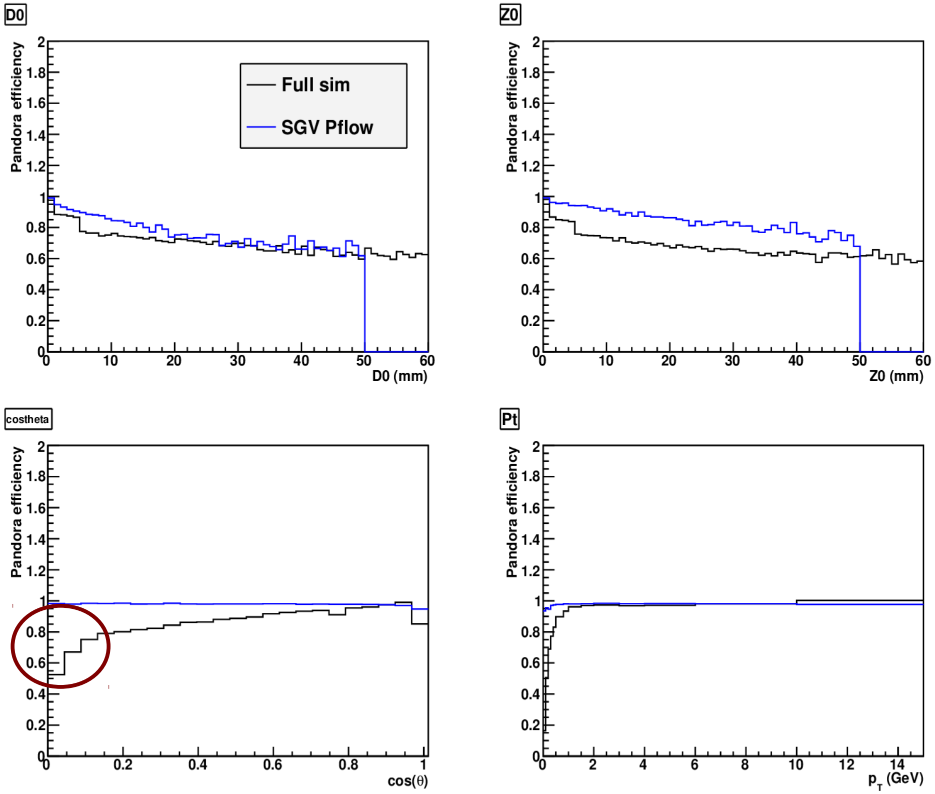
TrackMCTruthLink



Plot of weights indicating the contribution of a MC particle to a track and the inverse relation

Backup : Parametrizing Particle Flow Performance

- Pandora track selection :
  - Efficiency in full sim : tracks used by Pandora (associated to PFO) / tracks left after selection
  - Efficiency in SGV : tracks left after selection / tracks associated to PFO (by definition of SGV, all tracks are associated to a PFO)
- Tracking selection → good efficiency but in the central region not enough (due to curlers)
- In SGV : could add an efficiency for forming PFOs (not all tracks pass through the selection)



Backup : Comparison SGV Pflow / full simulation

- With Pflow
  - More double counting & less lost energy in SGV
  - Correlation E_{dc}/E_{lost} needs to be tuned for SGV